SECURITY CLASSIFICATION OF THIS PAGE

AD-A230 488

REPORT DOCUMENTATION PAGE

DTIC
ELECTE
JAN 09 1991
S D
B

a. SECURITY CLASSIFICATION AUTHORITY

b. DECLASSIFICATION / DOWNGRADING SCHEDULE

PERFORMING ORGANIZATION REPORT NUMBER(S)

1b. RESTRICTIVE MARKINGS

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release;
distribution unlimited.

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 1218

a. NAME OF PERFORMING ORGANIZATION
Anne Treisman
Department of Psychology
University of California, Berkeley 94720

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
Same as 8a

c. ADDRESS (City, State, and ZIP Code)
The Regents of the University of California
c/o Sponsored Projects Office, M-11 Wheeler
Berkeley, CA 94720

7b. ADDRESS (City, State, and ZIP Code)
Same as 8b

a. NAME OF FUNDING / SPONSORING ORGANIZATION
AFOSR/NL

8b. OFFICE SYMBOL (If applicable)

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR-87-0125

c. ADDRESS (City, State, and ZIP Code)
Building 410
Bolling AFB
Washington, DC 20332-6447

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2313	A4	

11. TITLE (Include Security Classification)
Visual information-processing in the perception of features and objects.

12. PERSONAL AUTHOR(S)
Anne Treisman

13a. TYPE OF REPORT
Final Technical

13b. TIME COVERED
FROM 1/1/87 TO 08/31/90

14. DATE OF REPORT (Year, Month, Day)
12/10/90

15. PAGE COUNT
13

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (~~Continue on reverse if necessary and identify by block number~~)
Features, Objects, Attention, Vision,

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Abstract.
The research explored various aspects of visual processing in the perception of features and objects. Studies of visual features tested the level of abstraction of the coding of orientation, the emergence of features in surface media defined by texture, motion or depth, and the role of similarity in preattentive representation of coarsely coded values. Object perception was explored in a series of studies requiring subjects to conjoin features, and testing the role played by spatially selective attention. The results suggested a modification of feature integration theory, allowing some preattentive control through feature maps. We also found evidence for object-specific tokens in the coding of dynamic displays of moving, changing objects. In studies of perceptual learning, extended practice effects were shown to be very specific to details of the practice situation. Finally, visual memory for arbitrary line patterns showed both surprising plasticity (long-lasting traces after a single presentation of 50 patterns) and surprising dissociations between implicit and explicit tests.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION

22a. NAME OF RESPONSIBLE INDIVIDUAL
Dr. Alfred R. Fregly

22b. TELEPHONE (Include Area Code)
(202) 767-5021

22c. OFFICE SYMBOL
NL

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted.
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

Final Technical Report: AFOSR grant # 87-0125
January 1987-August 1990.

AFOSR-TR- 1218

Background.

The research conducted on this grant concerned the visual coding of features and objects, the role played by attention, and in the fi ‑l two years some studies of perceptual learning and visual memory. The theoretical framework within which we explored a number of separate issues was the feature integration theory proposed by Treisman & Gelade (1980). The basic claims were that in the early stages of visual processing a number of elementary features are extracted in parallel in different, at least partly independent subsystems or modules. In order to specify how these features should be localized relative to one another and conjoined to specify objects, attention is focused seıially on each item or group of items in turn. Evidence consistent with this view included (1) the contrasting patterns of latencies in search for feature-defined and for conjunction-defined targets, suggesting a serial self-terminating scan for conjunction targets; (2) the occurrence of illusory conjunctions when attention is overloaded (Treisman & Schmidt, 1982), and (3) the much larger benefit in identification of conjunctions compared to features when attention is cued in adva..ce to the relevant location (Treisman, 1985). We also found that search for the absence of a feature (e.g. the only circle without a slash among circles with slashes) appears to be serial, as it should be if attention is needed to localize the only element which lacks a particular feature (Treisman & Souther, 1985). Search for the presence of a separable feature on the other hand is parallel. In this case the strategy could be to check the relevant feature detectors for the presence of the activity that uniquely characterizes the target. Prior localization is unnecessary.

The features linked to the currently attended location are entered into what we have called an "object file" - a temporary, episodic structure containing the information accruing about a particular object (Kahneman & Treisman, 1984). The object file is initially defined and subsequently addressed by its spatio-temporal location . Once the attended features have been assembled and entered into a shared object file, the contents are compared to stored descriptions in a recognition network and the likely identity, category and label can be found. The separation we proposed between object tokens (the object files) and object types (the stored representations in the recognition network) has important implications. It makes it easier to account for (1) the perception of repeated identical elements, (2) the perception of new arbitrary conjunctions of properties that have not previously been seen, and therefore have no stored representations to reactivate, and (3) our ability to update a single current object file when the object moves or changes without losing its perceptual continuity.

The new research supported by the grant from AFOSR can be classified under a number of different headings, as follows.

I Feature Analysis and Preattentive Processing.

The questions studied have concerned the nature of the visual coding that occurs early, automatically and in paral!el. We look for converging operations to tap these early levels of analysis. So far, we have used three different paradigms - search, texture segregation and apparent movement. We have also tried to probe the nature of the features extracted in these tasks, testing the level of abstraction at which orientation is defined, the possible equivalence of subjective and real contours, and of features of shape defined in different media (luminance, color, stereoscopic depth, differences in motion, etc.).

(1)Visual search and modularity.

Several experiments used visual search tasks with targets defined by simple features (orientation, color, size, and the presence of a gap in rectangular bars). The factors we varied were the number of different types of distractors present in any display and the number of different types of targets that were relevant on any trial. We argue that heterogeneity of the distractors on irrelevant dimensions should not affect search if the target is detected by analysis within a specialized module coding only the relevant target-defining feature. This is what we found for targets defined by color, orientation, and size. Conversely, we argue that if feature analysis is indeed modular, detection of a target might be slower when the relevant module is not specified in advance, so that subjects are forced to search for the odd one out. Again this is what we found (see pages 12-15 in Treisman, 1988). The results conflict with the hypothesis, proposed by Beck (1982), by Sagi and Julesz (1985), and implied by Marr (1982), that early stages of visual coding result in a single global representation pooling information about boundaries and discontinuities on all dimensions of variation.

(2)Apparent movement.

Apparent movement is seen when one or more elements are presented successively in different locations at the right temporal intervals. When more than one element is present, the perception of apparent movement requires that a match be made between elements in the first and second fields to determine which is seen to move where. This is known as the "correspondence problem" (Ullman, 1979). Since apparent motion is determined at short intervals and globally for a whole display, it is probably an early visual process, dependent on preattentive coding. It might therefore offer converging evidence for the psychological reality of particular perceptual features (Treisman, 1986). Ramachandran (1988) used this logic to show that shape-from-shading is available to be matched across successive displays and to generate apparent motion of a group of convex shapes against a background of concave shapes. We have adapted his method to test whether the correspondence required for apparent motion can be based on color, on orientation and on size. A subgroup of bars is embedded in a display containing two other types of bars that differ from them in color or size or orientation. The target group is shifted either as a whole, preserving all its spatial relations, or piecemeal (in different directions or distances), in a second display. The background items are shifted much less and in randomly selected directions, simply to eliminate cues from offsets and onsets. Subjects are asked to discriminate coherent from incoherent motion of the target group. The results so far suggest that both color and orientation can provide an input to the matching process that determines apparent motion, provided that they are highly discriminable. Size is less effective, although performance was better than chance, especially when the targets were larger than the background. When the target group was defined by a value intermediate between two types of distractors (medium-sized against large and small distractors, vertical against left and right-tilted distractors, gray against red and green distractors) performance was very poor for most subjects. Even here, however, a few subjects may be able to perform well above chance. This research is still in progress.

(3)Subjective contours as features?

Van der Heydt, Peterhans and Baumgartner (1984) have shown that single units in area V_2 of cat cortex respond to subjective contours. It seems possible, then, that subjective contours are coded automatically in early vision. Marcia Grabowecky and I are working on two tests of this hypothesis. (a) In a visual search task, we presented a target subjective triangle among "pacman" triples that do not create subjective triangles, and the converse - a pacman triple as target among subjective triangles, to see if the triangle would pop out while the non-triangle pacman triples did not. The results showed no pop-out and no search asymmetry with these stimuli. Both gave apparently serial search with about the same slope against display size. Informal observation suggested that not only focused attention but also visual fixation was necessary for the subjective contours to emerge. (b) Grabowecky is now testing whether the difficulty in seeing subjective contours off the fovea can be removed by scaling up the stimuli to equate acuity. If this succeeds, she will redo the search experiment to see if parallel processing is possible when acuity is high for the subjective triangles, but not for the pacman controls, or whether a difficulty still remains with the subjective contorus. If search is still serial, this would suggest that attention is needed to relate the three component shapes that create the illusory triangles and that the subjective contours emerge only once the occluding figure has been inferred. (c) We are looking for apparent movement of a subjective triangle between two pacman triples and testing whether the sensation of motion disappears when more stimuli are present and/or when eye movements are prevented. Again, the results so far suggest that the illusory triangle is seen to move when it receives attention and the eyes can follow it, but not otherwise. If these results hold up, they suggest that the coding of subjective contours is not automatic, but depends on attention. The cells in V_2 found by Van der Heydt may be part of a recurrent pathway with feedback from higher visual centers in the cortex.

(4)Coding of orientation.

Kathy O'Connell and I completed a series of experiments probing the nature of the representation formed of orientation at early stages of visual coding. It is known that a target line that differs in orientation from a set of background or distractor lines will "pop out" in a search task, allowing equally fast detection whatever the number of items in the display. We asked whether the orientation difference is coded in an analogue way (for example, by filters like the oriented receptive fields described by Hubel and Wiesel (1967), that sum the energy at a particular orientation), or whether it is coded more abstractly or symbolically, as suggested by Marr (1982). For example, would the same representation be formed for the orientation of a virtual line linking a pair of dots as for a solid line, or an edge? If so, would the direction of contrast need to be the same, both between lines or dot pairs and within dot pairs? We used a conjunction search paradigm to test whether shared orientations cause interference across different types of oriented stimuli. In a typical experiment, subjects might search for a line tilted right among dot pairs tilted right and lines tilted left. If the same orientation codes are activated by dot pairs and by lines, the target would be defined only by a conjunction of orientation (right) and "medium" (line rather than dots) and should therefore require serial checking with focused attention. This is what I found in an earlier study (Treisman, 1985). Kathy O'Connell and I extended the paradigm to test bicontrast dot pairs (one black and one white on a gray background) among mixed black and white lines. We found that orientation does not seem to be coded in parallel for bi-contrast dot pairs in the same way as it is for uni-contrast pairs. Although subjects could find the pair with the target orientation, the search process appeared to be serial. A paper on this research will soon be submitted.

We also ran an experiment following the same logic, testing whether lines and edges share the same orientation code. Here we found clear evidence that they do. Both conjunction conditions gave apparently serial search. In addition, the results suggested an asymmetry. The edge targets were found more slowly among lines of the same orientation than the reverse. This is consistent with physiological results: cells with asymmetric receptive fields ("edge detectors") will also respond to lines more than those with symmetric receptive fields will respond to edges. These results are consistent with the conclusion that parallel 'pop-out' for orientation targets is mediated by spatial luminance filters like the cells described by Hubel and Wiesel (1968). However, the results described in the next section suggest this is not a complete account of orientation coding.

(5)Different channels coding features of shape.

Together with Patrick Cavanagh and Martin Arguin, I completed a paper describing a series of experiments on search for targets differing in size or orientation, where the shapes (both targets and distractors) were defined by discontinuities in a number of other channels or media. We tested five different channels (previously studied in other perceptual tasks by Cavanagh): luminance, color, motion, stereoscopic depth and texture. We found parallel detection of size and orientation targets in every case except perhaps stereoscopic depth, where the discontinuities themselves were hardest to discriminate. We also found the same search asymmetry between vertical and tilted targets in the orientation domain across all the channels. The results suggest that the same vocabulary of shape-defining features may be extracted within a number of different media, (see Cavanagh, Arguin & Treisman, 1990).

(6) Feature similarity effects in search.

Duncan and Humphreys (1989) have recently proposed a "resemblance theory" of search, in which they claim that search latencies are completely determined by two measures of similarity: latencies are assumed to increase (a) with the similarity between target and distractors, and (b) with the dissimilarity between different distractors. They suggest that those two factors can also account for the difficulty of conjunction relative to feature targets in visual search. We attempted to test this claim by matching the similarity structure as closely as possible and comparing search for feature and for conjunction targets.

The stimuli were bars varying in color and orientation. Conjunction targets shared their color with one set of distractors and their orientation with another. Feature targets had a unique color and orientation, but each was more similar to the two distractor colors and orientations than those of the conjunction targets. Similarity was measured separately for each pair of relevant stimuli (both targets with distractors and distractors with distractors) by equating the search time for each feature in a background of the other. When the same distractors were paired to give a conjunction target, search was significantly slower than when they were paired to give a feature target, despite the fact that both similarity relations specified by Duncan and Humphreys were matched. We conclude that feature similarity cannot explain the difficulty of search for conjunctions (Treisman, 1991). The experiment revealed a further new result which is discussed in section II (2) below.

II The Perception of Conjunctions.

Feature integration theory proposed that once the various features present have been detected and grouped within separate modular maps, they are conjoined to form representations of the various objects in the scene by a process of serial scanning with focused attention.

1) Conjunction search.

One of these sources of evidence has recently been challenged, however. Nakayama (1986), Wolfe, Cave and Franzel (1989) and other investigators have found apparently parallel search for conjunction targets when their component features are highly discriminable colors, orientations, shapes, sizes or directions of motion. In order to try to understand the discrepancy between these results and my earlier findings, Sharon Sato and I completed a series of studies exploring conjunction search with conjunctions of highly discriminable features. We ran an experiment testing all possible pairings of highly discriminable values on the dimensions of color, size, orientation and direction of motion. As previously found by Nakayama (1986) and more recently by Wolfe, Cave and Franzel (1989), the search rates we obtained were substantially higher than in my earlier research, and in a few cases search seemed to reflect almost parallel coding. In addition, we found evidence consistent with the idea that each dimension contributed a constant to the slope of the search function.

I proposed a modification to my original feature integration theory to explain these results (Treisman, 1988). The suggestion is that spatial attention can be directed in two different ways: (1) As suggested in the earlier theory, it can scan locations serially through a selective "window" limited to one contiguous area at a time; (2) it can inhibit disparate locations in parallel on the basis of the non-target features they contain, perhaps by control through the separate feature maps I postulated as the output of parallel, preattentive processing. We ran further experiments testing the revised model: (1) We showed that a test of grouping salience correlated highly with the ease of conjunction search across the six targets we tested, as it should if inhibition from feature maps can help to articulate the visual field into separate sets of stimuli; (2) we showed that the rapid search rates depend on prior knowledge of the targets, as they should to allow inhibition based on non-target features; (3) we showed that search is faster when the conjunction target differs in two features from each distractor than when it differs only in one (cf Wolfe et al. 1989); (4) finally, we found that search is slower when the distractors are more heterogeneous, even when their average similarity to the target decreased. This is consistent with the idea that rapid search depends on inhibiting locations with non-target features rather than activating those containing target features. This research is described in Treisman & Sato (1990).

(2) Illusory conjunctions.

The feature-integration theory we proposed in 1980 and 1982 was an attempt to explain data from a number of different paradigms, not only from visual search. Some of the strongest support for the idea came from the occurrence of illusory conjunctions of features when attention is overloaded and serial processing is prevented. Given the flatter search functions obtained with highly discriminable features, it seemed important to test whether illusory conjunctions would still occur with similar simple and discriminable computer-generated stimuli as with the earlier tachistoscopic colored letters and shapes used by Treisman and Schmidt (1982). We confirmed that they do occur, about as frequently as with the earlier colored letter displays, with the dimensions of color, orientation and "medium" (i.e. outline, filled and broken rectangles).

Figure 1 shows a typical display. Immediately after the display, a mask, together with a bar-marker appeared, indicating which of the four stimuli should be reported, or whether the digits should be reported. By sampling the digits on separate trials from the colored bars, we reduced the possibility that decay in memory would cause conjunction errors. The digits were cued on 30% of the trials but were given much higher priority in the instructions, to ensure that attention would be spread across the display rather than focused on any single rectangle. In each case for each of the four features of the cued item, one incorrect value was present in the display - either once or three times - and one was not present.

Subjects reported an illusory conjunction on average on 25% of trials (this is the difference between the erroneous reports of a feature that was present in a non-cued location and a feature not present in the display). In the earlier Treisman and Schmidt experiment with colored letters the proportion of trials that gave illusory conjunctions of shape and color was 24%. The only dimension for which we found no clear evidence of illusory migrations in the present experiment was size: one reason might be that the size values were too hard to identify with the brief exposures we used. Subjects made errors on 47% of the sizes, compared to 34% of the media, 40% of the orientations and only 12% of the colors.

In this experiment we also asked a further question not previously tested: would the number of replications of a feature affect the probability that any one of them would migrate to form an illusory conjunction? My suggestion was that we conjoin features by attending to their locations. This allows access to all the features in the attended locations. When we do not attend , we still have information about which features are present, and about which locations are filled, but the links between "what" and "where" are not explicitly available. This account makes a prediction about what it is that migrates when an illusory conjunction is formed. The feature maps code only the presence and perhaps the amount of each feature, without locating any instances of it and without indviduating the occurrences. Without attention they might tell us for example that there is a great deal of red and a little green, and that there is some horizontal and some vertical. Meanwhile the location map might tell us that there are several things here and a few there, but it would not tell us what those things are. So what migrates should not be the individual instances of a feature but simply its presence. Errors should therefore take the form of mislocations of where a particular feature is. This means that the number of elements that are red should not directly predict the probability of an illusory conjunction containing red. All that should matter would be the detection of the presence of red somewhere in the display: the type rather than the tokens.

This research is still in progress. So far the results show ratios that are significantly less than 3 to 1 for illusory conjunctions involving the migration of a feature replicated three times in the display and one presented only once. The ratio is also significantly greater than 1 to 1, averaging 1.7. The difference could, however, reflect the greater probability of detecting the presence of the relevant feature given three instances than one. A control experiment is needed to see whether detection rates match this 1.7 to 1 ratio, and also to see how accurately the number of instances can be estimated when the size of each or the total quantity of the relevant feature is controlled.

3) Independence of identity and location for features.

Another way of looking at the relations between the location map and the feature map is to ask for report of both what and where and to see whether and under what conditions the two are inter-dependent or independent. Treisman and Gelade had found that for conjunctions "what" and "where" are completely interdependent. No identity information was available when the localization was incorrect and no localization was possible when the identity was wrongly reported. However we had found considerable independence of "what" from "where" when the targets were simple features.

Recently, I have tested this result further, partly because Johnston and Pashler found a possible artefact in our method and failed to replicate our result for distant location errors. We also extended the experiment by using both homogeneous and heterogeneous distractors. Subjects looked for any of 12 possible targets among the homogeneous distractors, essentially detecting the odd one out. They looked for any of 3 possible targets - blue, tilted or broken texture - among the heterogeneous distractors. We are finding rather different results. In both cases, subjects get many features correct given a location error. Corrected for guessing, they average around 30% to 40% correct. With heterogeneous distractors, they also perform significantly above chance in reporting identity even when a non-adjacent location is reported (around 20%). The results support my previous conclusion. When the distractors are homogeneous, identity information is still better than chance when a non-adjacent location is reported, but the accuracy is much lower (only about 7%, corrected for guessing). Looking at correct localizations given that the identity was wrong, subjects again perform well above chance with homogeneous distractors, (around 32%), but no information is available about the target location when its identity is incorrect and the distractors are heterogeneous. To make sense of these results, I think we have to say that the location map signals feature differences although it does not specify what the features are. It allows us to localize the one unique feature without knowing in what way it is unique. This would allows us to locate boundaries between distinct areas; it would help in segregating figure from ground, but it would not distinguish the nature of the discontinuity. In the heterogeneous displays, there are many discontinuities, so the target cannot be localized unless its identity is also known.

4)Iconic memory.

Marcia Grabowecky completed her M.A. thesis on three studies of iconic memory for conjunctions of color and shape. She presented colored letters at 8 locations in a circular array and cued which one should be reported, by presenting a white dot just outside one of the letter locations, at different intervals relative to the onset of the display (from 0 to 1000 msec.). In one condition, subjects reported only the color of the cued letter; in another condition they reported only its shape; and finally in a third condition they reported both its color and its shape. If conjunction information is present at any stage, we predicted that the probability of getting both correct on conjunction report trials should exceed the product of the probabilities of getting the color correct and of getting the shape correct. On the other hand if color and shape are registered independently, the conjunction information might take time to emerge, or might never emerge if attention is not focused on the correct item in time (i.e. if the cue is presented late relative to the display). Marcia found that the results fitted the independence prediction at all intervals tested. The result is consistent with the prediction from feature integration theory, that feature are initially registered independently and that they are combined only through focused attention. It rules out the alternative possibility that conjunction information (e.g. "red T-ness" or "Q-like blueness") is initially present but is rapidly lost unless attention is focused to maintain it.

5) Similarity and within-dimension conjunctions.

In the experiments on similarity effects on search described in section I(6) above, we found a large difference in our results, over and above the effects of feature vs. conjunction targets. A target defined by "standard values" (blue and vertical) was found much more easily, both in feature and in conjunction search, than targets defined by non-standard values (purple or turquoise tilted 27°). I suggested that coarse coding may be used at the preattentive level, giving direct access only to a few standard values on each dimension (e.g. red, green, blue and yellow for color; vertical, horizontal and left and right diagonals for orientation). Other values would then be coded by ratios of activity in pairs of these standard populations of detectors, - in other words they would be within-dimension conjunctions. Search for a purple target, coded as activity in both blue and red detectors, would require focused attention to conjoin its features when presented among distractors with the same standard values - blue and red bars.

Converging evidence for this hypothesis was obtained from a test for illusory conjunctions when the same search displays were presented briefly and followed by a mask. Subjects reported a substantial number of illusory targets in both conjunction and "feature" displays for the non-standard targets. Similarity may pose problems at least in part because the coarse coding of features creates conjunction problems when stimuli in the same display share the same underlying components (see Treisman, 1991).

6) Conjunctions within dimensions: Spatial separation.

Another way in which conjunctions of features can be formed within a single dimension is to present pairs of separate values (e.g. two colors or two orientations) in various spatial configurations. Wolfe, Yu, Stewart, Shorter, Friedman-Hill & Cave (1990) compared search for between and within-dimension conjunctions and claimed that whereas search for both between-dimension spatially integrated conjunctions (e.g. a red C among green C's and red I's) and between-dimension spatially separated conjunctions (e.g. a grey C with an intersecting red bar among red I's with gray bars) could be fast or parallel when the features are sufficiently discriminable (see section II(1) above), search for within-dimension conjunctions was always slow and serial (e.g. finding either a red C with a green bar or a green C with a red bar among red C's with blue bars, green C's with blue bars, grey C's with red bars and grey C's with green bars required at least 40 ms. to check each item).

Beena Khurana in my lab has run experiments which suggest that the difficulty may be in conjoining spatially separate elements rather than in conjoining values within a single dimension. She has run several experiments in which the spatially separated color-color targets are no harder than carefully matched spatially separated color-shape or color-orientation targets. She is testing possible uncontrolled factors that may have made the between-dimension conjunctions too easy in the experiment by Wolfe et al. If correct, her conclusion that the critical factor is spatial separation of the features rather than the fact that they are values on the same dimension has important implications for the theory.

III "Object Files" and the Integration of Information.

1) Moving visual objects and the "reviewing effect".

With Daniel Kahneman, I resumed a line of research that we had pursued some years earlier, exploring the perception of moving, changing objects and collecting evidence for object-specific representations. In a typical experiment, we presented two letters in two separate frames in a "preview" display, followed by motion of the empty frames to two new locations, and finally a single letter would appear in one of the two frames, to be named by the subject. We found a priming or "reviewing" benefit in naming the final letter when it matched an earlier letter, but only if the earlier letter had appeared in the same frame (in its previous location). We had run a number of experiments, exploring the temporal and spatial conditions in which this object-specific reviewing benefit was obtained, and extending the paradigm to look at the object-specific integration of features (the two lines of a plus). In the past two years, we returned to this project and ran some further experiments.

(1) We compared the reviewing benefit with stationary objects and showed that it could not be attributed to iconic memory, that it was affected by load (the number of displayed objects), and that the limit was set by the number of tokens in the display (instances of letters, whether or not

they were repeated) rather than types (different letter identities). The limit to the number of tokens in which letter and frame were integrated seemed to be between 2 and 3. We also ran many experiments with moving objects trying to distinguish the level at which information accrues - for instance is the priming case-specific , or letter-specific, or is there accumulation of evidence also for arbitrary categories that share a response. We also ran various control experiments to check our earlier findings and clarify their interpretation. We wrote and submitted (in August, 1990) the first of two or three papers reporting this project (Kahneman, Treisman & Gibbs, 1990, enclosed).

2) Cross-modal objects.

Meg Wilson has tried to extend the object file idea to predict cross-model integration of information. The task was to judge whether a pre-cued visual object was flashed once or twice, comparing performance when an auditory signal either agreed or disagreed in number (one or two tones). The manipulation was to make the auditory signal appear to come either from the same visual object or from another visual object. The two were either spatially or temporally separated , and there was a correlation of visual shape and sound (low tone with a fat shape and high tone with a thin one). So far, she has obtained a small but significant effect with spatial separation but not with temporal separation. Subjects were more influenced by the tone in judging the visual number when the two appeared to occupy the same spatial location, but there was no significant capture when she manipulated timing.

3) Attention span with moving objects.

Together with Meg Wilson, I have begun some research in a paradigm developed by Pylyshyn and Storm (1989) exploring subjects' ability to keep track of several randomly moving, otherwise identical objects. Pylyshyn and Storm found that subjects could keep track of up to 4 among 8 identical randomly moving shapes, so that when probed with one of the eight they could distinguish with about 90% accuracy whether it was one of their assigned targets or a distractor. This is an interesting finding in relation to the concept of "object files" that Kahneman and I proposed (1984); it suggests that once four separate object files have been set up, they can be maintained without devoting focused attention to each of the four. Pylyshyn (1988) has proposed the notion of a limited number of "FINSTs" or indices, which can preattentively maintain the spatial addresses of visual tokens as they or the observer move.

We decided to test how far the task really is independent of attention, by seeing whether it is disrupted by a concurrent attention-demanding task. We spent some time developing a task which would not also require visual fixation, (since we have no equipment to measure eye movements, and we wanted to leave subjects free to fixate where they wished). The attention task we developed involves monitoring a changing sequence of colors and textures in a broad frame (around the foveal area in which the objects move) in order to detect the occurrence of a particular conjunction of color and texture (eg. red vertical stripes) in a rapid sequence of other combinations of color and texture. At any one time, only one color and one texture are present, so there is no need to localize or fixate the stimuli in space. Instead they must be conjoined repeatedly for each brief interval of time. We compared subjects' ability to track the moving objects without a concurrent task, with a conjunction-monitoring task (e.g. watch the frame for red stripes in a sequence containing red spots and green stripes), and with a feature-monitoring task (e.g. watch the frame for blue or plaid in the same sequence of red spots and green stripes). We found a significant decrement in tracking when subjects also monitored the border, and the conjunction task suffered more than the feature task when combined with tracking the moving elements, suggesting an involvment of attention. There was also a significant decrement when subjects tracked 4 rather than 3 elements. However, we found no interaction between these two putative ways of increasing the load. We hypothesized that subjects attempt to form a mental representation of a global figure consisting of virtual lines joining the target elements, and to watch it distorting as the elements move. There would then be only one attended oject - a triangle or a quadrilateral - whether 3 or 4 dots were tracked. Yantis (1990) has since tested this hypothesis and obtained some supporting evidence. It remains puzzling, however, that our conjunction-monitoring task did not interfere more than our feature-monitoring task with maintaining attention to the global figure. Subjects may simply have given priority to the tracking task, producing the greater drop in accuracy with conjunction manitoring.

We plan to integrate the stimuli for the concurrent task with those for the moving object task, to see if it is easier to divide attention when it is the moving objects that change color and texture than when it is the surrounding frame. The experiment may throw light on the mechanisms of attention as well as on those involved in tracking and maintaining representations of moving object tokens. Attention may be limited in two different ways: (1) the number of spatially distributed objects it can encompass at any one time; (2) the number of tasks that can be performed at once, even on the same objects. The kind of attention I have studied within the framework of feature integration theory has been the spatially selective "window" that I claim specifies which features belong together. It is not clear that moving a 'mental' window around a scene need compete with temporally serial checking of a sequence of color-texture pairs. At most one object file (or FINST in Pylyshyn's mode) would be taken up by the frame task. No extra files or FINSTS would be involved in the version in which the moving objects themselves change color and texture. The results may indicate, therefore, whether attention capacity is required to conjoin features as well as to select spatially which features to conjoin.

IV Perceptual Learning and Visual Memory.

Once focused attention has established an integrated representation for a complex combination of parts or properties, the representation presumably becomes part of our visual memory. It is of interest to study how it persists or changes over time and how it affects the perceptual coding of later presentations of the same pattern.

1)Perceptual learning.

Alfred Vieira and I have explored the effects of prolonged practice at search for conjunctions of shape elements on other measures of visual processing with the same shapes. We completed one study of automatization in letter search with four subjects who each had 16 sessions of practice searching for three arbitrarily selected letters (EXR or TVQ) in displays of 1, 2 and 4 other letter distractors. We compared their performance before and after practice, both on the practiced targets and on the control targets (The practiced targets for 2 subjects were the control targets for the other 2, and vice versa). The tests we used were texture segregation, conjunction search and identification in contexts that might generate illusory conjunctions, perception of words containing the target letters, and target localization. The results suggest that the perceptual learning that progressively speeds search and reduces the slope of search functions against display size is highly specific to the particular search task. There is little change in subjects' ability to detect boundaries between an area containing the target letters and an area containing distractors in a texture segregation task; there is little decrease in the dependence of identification on localization and search improvements with upper case letters generalize only partly to search for the same letters in lower case.

The test for illusory conjunctions gave an unexpected result before practice: we found no evidence for illusory conjunctions for parts of shapes. For example, subjects did not form illusory R's from P's and Q's or E's from F's and L's. We were therefore unable to test whether these illusory conjunction errors decrease with automatization. The absence of illusory interchanges suggests that letters may actually be more integrally coded than we had thought, even before practice at search begins (see Treisman & Souther, 1986, for other evidence consistent with this conclusion).

The word perception task was designed to test whether automatic detection of individual letters makes it more difficult to read whole words that contain those letters. The subjects run by Schneider and Shiffrin (1977) in multiple search sessions complained that practice at search made it difficult to read the newspaper because all they could see were the targets they had learned to detect. We found, however, no difference at all in the speed of lexical decision between words containing the target letters and words containing control letters.

We conducted another study of search automatization using more complex arbitrary, meaningless shapes (six-line figures in a 3 x 3 dot matrix). Again, we obtained a very large increase in search rates over 16 sessions of practice. We then explored the degree of transfer to a large number of other perceptual tasks, including mental rotation, perception of apparent motion, iconic memory, threshold in same-different matching, finding parts in wholes, and ratings of clarity, goodness and likability. Again, we found remarkably little transfer. In an attempt to find out what was learned, we also looked at transfer to variations within the search task itself. We found a striking specificity to the particular learned targets, less to the learned distractors, and even a significant decrement when we changed irrelevant aspects of the display, such as the direction of contrast (black on green vs. green on black) or the spatial configuration of the displays. The results are consistent with an account of automatization in terms of the accumulation of specific exemplars of previous trials in memory (Logan, 1988). A paper describing this study was presented at the meeting of the Psychonomic Society in November, 1988 (copy enclosed), adn we are preparing a paper to submit for publication.

2) Visual memory; explicit and implicit measures.

Gail Musen and I ran one experiment together and she ran two further experiments on visual memory for nonsense figures (5-line versions of those described in section IV (1)). The three experiments together constituted her Ph.D. dissertation; the first was published this year (Musen & Treisman, 1990, enclosed).

Most work on visual memory has used either words or pictures of familiar objects. Explicit memory is typically measured by recall or recognition, whereas implicit memory is measured by priming of performance in a non-memory task such as perception near threshold or word-fragment completion (Tulving, Schacter and Stark, 1982). If performance is better for previously studied items, this is taken to reflect memory without awareness of the earlier presentation. The two measures have been shown to be independent with familiar verbal stimuli. The standard explanation has been that priming reflects the persisting activation of pre-existing representations or nodes in a semantic memory network, whereas recall and recognition depend on separate episodic memory traces.

In our experiment, we found clear perceptual priming after a single presentation of 50 of these novel visual patterns. The task was to draw each pattern immediately after it was briefly flashed and followed by a mask. The priming measure was the difference in the number correct for previously studied patterns and for new ones. The priming did not increase much with four further presentations or decrease much with the passage of time (up to a month), whereas recognition memory for the same patterns showed substantial effects of both. Moreover, performance was stochastically independent for the same patterns in the recognition and in the priming tasks. We also directly tested episodic memory for the patterns by showing half of the previously studied patterns one more time and seeing whether they differed from the other half by either explicit or implicit memory measures. We found excellent recognition performance in the task of discriminating the re-presented items from the others, and no difference at all in the priming they produced.

In the two further experiments for her dissertation, Musen showed a dramatic effect of verbal coding on recognition memory and no effect on priming; the second experiment showed somewhat reduced priming when the study time was reduced from 10 seconds per pattern to 1 second, but a much greater reduction in recognition memory.

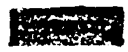
These memory dissociations with previously unfamiliar stimuli pose a problem for the account in terms of the separation of episodic from semantic memory, since no pre-existing representations of our line patterns were available to be reactivated and mediate priming. On the other hand, the stochastic independence and functional dissociations pose problems for an account that suggests that the same episodic memory traces mediate both tasks. Perhaps a new representation of a "type" can be set up in a single presentation; later recurrences are then natched to this type to facilitate perceptual identification, but also lay down separate traces for ach token of the type, to mediate later explicit memory tasks, like recall, recognition or amiliarity judgments.

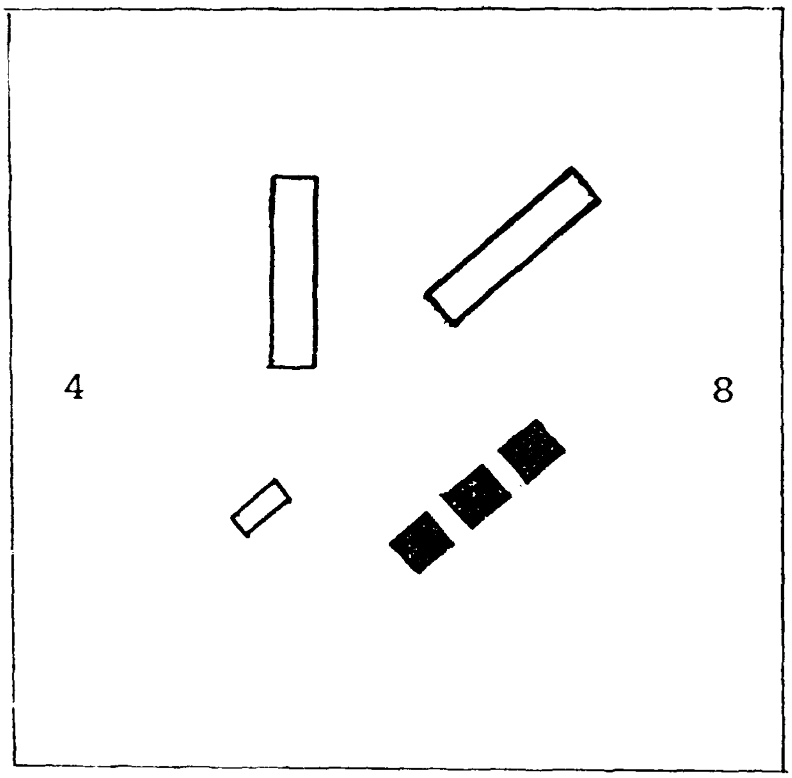
Gail Musen has now moved to a post-doc. with Larry Squire at UCSD, where she is testing mnesic patients in the same task, to see if they show the same selective loss of explicit but not mplicit memory with visual patterns as they do with verbal stimuli.

Figure 6

Sample display to test the occurrence of illusory conjunctions.

Values not shown
(red, horizontal,
filled, medium-sized)





References.

Beck, J. (1982) Textural segmentation. In J. Beck (Ed.), Organization and representation in perception. Hillsdale, NJ: Erlbaum.

Cavanagh, P. (1987). Reconstructing the third dimension: Interactions between color, texture, motion, binocular disparity and shape. Computer Vision, Graphics and Image Processing, 37, 171-195.

Hubel, D.H., & Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. Journal of Physiology, 195, 215-243.

Julesz, B. (1984). Toward an axiomatic theory of preattentive vision. In G.M. Edelman, W.E. Gall and W.M. Cowan (Eds.). Dynamic aspects of neocortical function. Neuroscience Research Foundation.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. To appear in R. Parasuraman and J. Beatty (Eds.), Varieties of attention. New York, Academic Press,

Kanwisher, N. (1987). Repetition blindness: Type recognition without token individuation. Cognition, 27, 117-143.

Logan, G. (1988). Toward an instance theory of automatization. Psychological Review, 95(4), 492-527.

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. San Francisco: W.H. Freeman & Co.

Nakayama, K. & Silverman, G. H. (1986). Serial and parallel encoding of visual feature conjunctions. Investigative Ophthalmology and Visual Science.

Nakayama, K., & Silverman, G.M. (1986). Serial and parallel processing of visual feature conjunctions. Nature, 320, 264-5.

Pylyshyn, Z.W. (1988). Here and there in the visual field. In Z.W. Pylyshyn (Ed.), Computational processes in human vision: An interdisciplinary perspective. Norwood, NJ: Ablex.

Pylyshyn, Z.W. & Storm, R.W. (1989). Tracking of multiple independent targets: Evidence for a parallel tracking mechanism. Spatial Vision. (in press).

Ramachandran, V.S. (1988). Perception of shape from shading. Nature, 331, 163-166.

Schneider, W., & Shiffrin, R.M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Review, 84,1-66.

Treisman, A. (1985). Preattentive processing in vision. Computer Vision, Graphics, and Image Processing, 31, 156-177.

Treisman, A. (1986) Properties, parts and objects. In K. Boff, L. Kaufman, & J. Thomas (Eds.), Handbook of perception and human performance. Volume 2, Cognitive processes and performance. New York: Wiley.

Treisman, A. (1988) Features and objects: the fourteenth memorial Bartlett lecture. Quarterly Journal of Experimental Psychology, in press.

Treisman, A., & Gelade, G. (1980) A feature integration theory of attention. Cognitive Psychology, 12, 97-136.

Treisman, A., & Souther, J. (1985) Search asymmetry: A diagnostic for preattentive processing of separable features. Journal of Experimental Psychology: General, 114, 285-310.

Treisman, A. & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. Psychology Review, 95(1), 15-48.

Treisman, A., & Schmidt, N. (1982). Illusory conjunctions in the perception of objects. Cognitive Psychology, 14, 107-141.

Tulving, E., Schacter, D.L. & Stark, H. (1982). Priming effects of word-fragment completion are independent of recognition memory. Journal of Experimental Psychology: Learning, Memory and Cognition, 8, 336-442.

Ullman, S. (1979) The interpretation of visual motion. Cambridge, Mass.: M.I.T. Press,.

Von der Heydt, R., Peterhans, E. & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. Science, 224, 1260-1262.

Publications during the period supported by the AFOSR grant.

Treisman, A., & Gormican, S., 1988. Feature analysis in early vision: Evidence from search asymmetries. Psychological Review, 95, 15-48.

Treisman, A., 1988. Features and object: The Fourteenth Bartlett Memorial Lecture. Quarterly Journal of Experimental Psychology, 1988, 40A, (2) 201-237.

Cavanagh, P., Arguin, M. and Treisman, A., 1990. Effects of stimulus domain on visual search for orientation and size features. Journal of Experimental Psychology: Human Perception and Performance, 16, 479-491.

Musen, G. & Treisman, A., 1990. Implicit and explicit memory for visual patterns. Journal of Experimental Psychology: Learning, Memory and Cognition, 16, 127-137.

Treisman, A., 1990. Search and the detection and integration of features. Behavioral and Brain Sciences, 13, 454-455.

Treisman, A. 1990. Variations on the theme of feature integration: Reply to Navon (1990). Psychological Review, 97, 460-463.

Treisman, A., Cavanagh, P., Fischer, B., Ramachandran, V. and Van der Heydt, R., 1990. Form perception and attention: striate cortex and beyond. Spillman, L. and Werver, J. (Eds.) Visual Perception: The Neurophysiological Foundations, New York: Academic Press.

Treisman, A. & Sato, S., 1990. Conjunction search revisited. Journal of Experimental Psychology: Human Perception and Performance, 16, 459-478.

Papers in press.

Treisman, A. 1990. Visual coding of features and object: some evidence from behavioral studies. Proceedings of NRC Symposium: Frontiers of Visual Science, in press.

Treisman, A. 1991. Search, similarity and the integration of features within and between dimensions. Journal of Experimental Psychology: Human Perception and Performance.

Treisman, A. 1991. Automaticity and preattentive processing. Symposium on "Varieties of Automaticity" to be published in special issue of American Journal of Psychology.

Treisman, A. 1991. Representing visual objects: A discussion of session on "Visual information processing." Attention and Performance XIV, D. Meyer and S. Komblum (Eds.), in press.

Kahneman, D. Treisman A. & Gibbs, B., 1991. The reviewing of object files: Object-specific integration of information. Cognitive Psychology.

Awards and Honors.

June, 1989. Elected a Fellow of the Royal Society, London.

March, 1990. Howard Crosby Warren Medal of the Society of Experimental Psychologists.

Invited lectures and conference papers.

January 1987. Fourteenth annual Sir Frederic Bartlett Memorial Lecture, to Experimental Psychology Society, London, England.
Two Papers to Bat Sheva Seminar on 'Selective attention in sensory processing' in Jerusalem, Israel.

February 1987. Lecture to Cognitive Science program, University of California, Berkeley.
Colloquium to Psychology Department, Stanford University.
Talk to Smith Kettlewell Institute, San Francisco.

March 1987. Paper to National Research Council Committee on Vision Symposium on 'Frontiers in Visual Science,' Washington, D.C.

May 1987. The Paul Fitts Memorial Lectures at the University of Michigan on 'Attention, features and objects,' jointly with Daniel Kahneman, to be published as a book.

June 1987. Paper to conference on 'Les Approches de la Cognition' at Cérisy-la-Salle, France, to be published in a book.

July 1987. Paper to conference on the 'Neurophysiological foundations of visual perception' in Badenweiler, Germany.

November 1987. Invited paper in Presidential Symposium on Attention at the Neurosciences conference in New Orleans.

January 1988. Colloquium to University of California, Santa Cruz.

March 1988. Paper to Howard Hughes conference on Vision, Miami.
Invited Paper to American Association for Artificial Intelligence, Stanford.

April 1988. Invited address to Western Psychological Association, San Francisco.

June 1988. Invited lecture to McDonnell Summer Institute on Cognitive Neuroscience, Harvard University.

September 1988. Talk to AFOSR Workshop on attention, Colorado Springs.

October 1988. Paper to Canadian Institute of Advanced Research Annual Meeting, Whistler, British Columbia, Canada.

November 1988. Vieira and Treisman. Paper to Psychonomic Society, Chicago. Colloquium at U.C. San Diego.

January 1989. Invited lecture to Cognitive Neurosciences Program at University of Oregon.

May 1989. Colloquium at University of California, Riverside.

June 1989. Discussant at conference to honor Wendell Garner, Yale University.

January 1990. Invited paper at National Research Council Committee on Vision, Conference on Attention and Search, Irvine, California.

April 1990. Invited paper to Conference on 'Recent Advances in the Analysis of Attention', Eugene, Oregon.

April 1990. Colloquium and seminar at Washington University, St. Louis.

May 1990. Paper on 'Search and Similarity' at ARVO, Sarasota, Florida.

May 1990. Colloquium in Cognitive Science and Distinguished Lecture at U.C.L.A.

June 1990. Invited address to the American Psychological Society.

July 1990. Invited address to Cognitive Science Society, Cambridge, MASS.

July 1990. Discussant at 'Attention and Performance XIV, Ann Arbor, Michigan.

Personnel whose research was supported wholly or partly by the grant.

Anne Treisman, Principal Investigator
Sharon Sato, Technical assistant, 50% time, January 1987-July 1989
Amy Hayes, Technical assistant, 50% time, August 1989-August 1990
Ephram Cohen, Programmer, 15% time
Stephanie Muth, Sarah Preisler, Nicola Walters, Sherlyn Jimenez, Kathleen Miszuk, Julia Simovsky - Clerical assistants (15% time, in succesive periods).

Graduate students:

Gail Musen, 1989 Ph.D.
Alfred Vieira, advanced to condidacy, 1990.
Marcia Grabowecky
Beena Khurana
Kathy O'Connell
Brett DeSchepper
Elizabeth Beiring
Meg Wilson